

# Theoretical predictions for vehicular headways and their clusters

Milan Krbálek and Katarína Kittanová

*Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague,  
Prague – Czech Republic*

---

## Abstract

This article presents a derivation of analytical predictions for steady-state distributions of netto time gaps among clusters of vehicles moving inside a traffic stream. Using the thermodynamic socio-physical traffic model with short-ranged repulsion between particles (originally introduced in [Physica A **333** (2004) 370]) we firstly derive the time-clearance distribution in the model. Consecutively, the statistical distributions for the so-called time multi-clearances are calculated by means of theory of functional convolutions. Moreover, all the theoretical surmises used during the above-mentioned calculations are proved by the statistical analysis of traffic data. The mathematical predictions acquired in this paper are thoroughly compared with relevant empirical quantities and viewed in the context of traffic theory.

*Keywords:* vehicular traffic, headway distribution, socio-physical traffic model

*PACS:* 05.40.-a, 89.40.-a, 05.45.-a

---

## 1. Introduction

The most of scientific works dealing with traffic theory do not omit to discuss certain regularities occurring in traffic microstructure. Explorations of traffic micro-quantities (i.e. quantities belonging to an individual vehicle) and efforts to predict their statistics are as old as traffic research itself. Virtually, all important reviews on traffic science (for example [1, 2, 3], or [4]) try to explain at least some basic knowledge on statistical distributions of time-intervals or distance gaps among moving cars. Although the theoretical prediction of time-evolution for headway-distributions is still extremely vague, some partial achievements have been reached in the last years (for example [5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16],

or [17]). We try to pick up the threads of those results and get closer to the heart of the matter.

This contribution is focused predominantly on the time-evolution of clear time intervals (so-called *time-clearances*) between succeeding cars passing a given point (a traffic detector, typically) located at an expressway. Investigations the traffic clearances are significantly advantageous (contrary to explorations of distance gaps) because of their direct measurability. Indeed, the most of traffic detectors gauges the time of vehicle's passage directly, which means that those data do not show any systematic or mediated errors (provided the detector is not damaged). Except the time-clearances we concentrate our attention on the so-called *multi-clearances*, i.e. cumulated time-clearances among  $n$  consecutive cars as well. Again, such a quantity is (similarly to time-clearance itself) directly measurable, which opens a possibility for detailed statistical analyses of relevant multi-clearance evolution (i.e. investigation of multi-clearance distributions with respect to the location of traffic systems in the fundamental diagram).

In fact, the empirical and/or theoretical investigations of traffic multi-headways are not sporadic in the physics of traffic. Some of the previous scientific works, e.g. [18, 19, 20, 21], or [22], briefly analyze the time multi-clearance distributions (predominantly from the empirical point of view) or their statistical variances. The reasons for investigations of such a type are obvious, since deeper comprehension of changes in traffic microstructure will provide a more thorough insight into the convoluted traffic interactions.

## 2. Preliminary evaluation of empirical observations

The vehicular-data records analyzed in this article are typically of the following types. The set

$$T_\ell^{(\text{in})} = \{\tau_{k\ell}^{(\text{in})} \in \mathbb{R}^+ \mid k = 1, 2, \dots, N\}, \quad (\ell \in \{0, 1, 2\})$$

includes the chronologically-ordered times when the front bumper of  $k$ th car has intersected the detector-line (located at  $\ell$ th lane of an expressway). The respective times  $\tau_{k\ell}^{(\text{out})}$  when the rear bumper of  $k$ th car has intersected the detector-line are summarized in the set  $T_\ell^{(\text{out})}$ . Analogously, the set of velocities of individual vehicles recorded by the detector is

$$V_\ell = \{v_{k\ell} \in \mathbb{R}^+ \mid k = 0, 1, 2, \dots, N\}, \quad (\ell \in \{0, 1, 2\}).$$

The associated lengths of vehicles are denoted as  $d_{k\ell}$  and summarized in the set  $D_\ell$ . We remark that the indexes  $\ell = 0, 1, 2$  correspond to the slow, main, and overtaking lanes of a freeway, respectively. The slow lane is intended predominantly for long vehicles (trucks or lorries) and therefore the respective data will not be considered for this research. On contrary, the drivers use the third lane ( $\ell = 2$ ) particularly if they are overtaking (and after overtaking-manoeuve they turn back in the main lane) or if they are significantly faster than other cars. All of the above-mentioned quantities are (for purposes of this article) considered to be primary, which means that they are directly detectable by the traffic detectors.

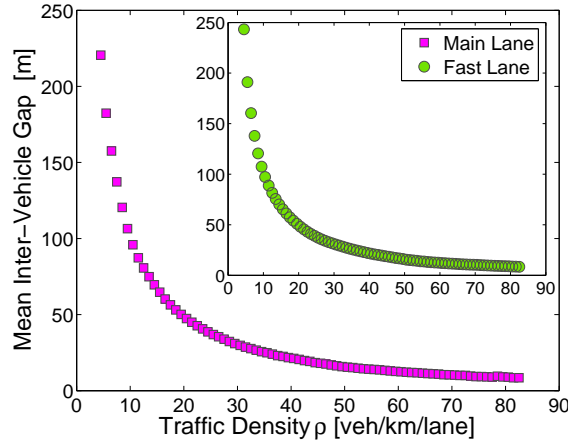


Figure 1: Sample-mean clearance of subsequent vehicles. The average value of real (non-scaled) space-clearances among cars is decreasing function of traffic density. The difference between lanes is fractional.

Except the primary traffic quantities we now introduce some important secondary quantities, whose values are obtained vicariously, i.e. they are not included in the traffic-detector's records. Among the secondary traffic micro-quantities they use to be accented *the time headways*

$$z_{k\ell} := \tau_{k\ell}^{(\text{in})} - \tau_{(k-1),\ell}^{(\text{in})}$$

and *time clearances*

$$t_{k\ell} := \tau_{k\ell}^{(\text{in})} - \tau_{(k-1),\ell}^{(\text{out})}.$$

Although both of them are not explicitly included in traffic-data files, their values are not burdened with any additional error. Above that, brutto space-gaps

between successive vehicles (usually called as *the distance headways*) are traditionally approximated by the relation  $s_{k\ell} := v_{k\ell} z_{k\ell}$  that presupposes the constant velocity  $v(\tau) = v_{k\ell}$  during a time period when  $\tau \in [\tau_{(k-1),\ell}^{(\text{in})}, \tau_{k\ell}^{(\text{in})}]$ . As well known, such a precondition is questionable, especially in the region of small traffic densities where the time headways are too large. However, the influence of a possible error is of marginal importance, as apparent from the fact that the headway distributions analyzed in small-density regions do not show any noticeable deviation from exponential distribution expected for infrequent events (see [11, 2], or [6]). Analogously, *the distance-clearance* is calculated via  $r_{k\ell} := v_{k\ell} t_{k\ell}$  and represents the estimated netto distance between  $k$ th car and its predecessor. Contrary to the time headways (clearances) the distance headways (clearances) are burdened by the systematic error that is discussed above. Such a error devalues the knowledge on distance headway distributions and their evolution.

Denoting the sampling size by  $m$  (in this study there is consistently considered  $m = 50$ ) and number of data-sets by  $M_\ell$  (which therefore implies that  $M_\ell = \lfloor N_\ell/m \rfloor$ ), where  $N_\ell = \text{card}\{\tau_{k\ell}^{(\text{in})} | \ell \text{ is fixed}\}$ , one acquires the main data-samples

$$S_j^{(\text{main})} = \{(\tau_{k\ell}^{(\text{in})}, \tau_{k\ell}^{(\text{out})}, v_{k\ell}, d_{k\ell}) \in T_1^{(\text{in})} \times T_1^{(\text{out})} \times V_1 \times D_1 | \\ k = (j-1)m + 1, (j-1)m + 2, \dots, jm \wedge \ell = 1\}, \quad (1)$$

and the secondary data-samples

$$S_i^{(\text{fast})} = \{(\tau_{k\ell}^{(\text{in})}, \tau_{k\ell}^{(\text{out})}, v_{k\ell}, d_{k\ell}) \in T_2^{(\text{in})} \times T_2^{(\text{out})} \times V_2 \times D_2 | \\ k = (i-1)m + 1, (i-1)m + 2, \dots, im \wedge \ell = 2\},$$

where  $j = 1, 2, \dots, M_1$  and  $i = 1, 2, \dots, M_2$ . For each data-sample  $S_j^{(\text{main})}$  (or  $S_i^{(\text{fast})}$  alternatively) we calculate the local flux

$$J_j = \frac{m}{\tau_{jm,\ell}^{(\text{out})} - \tau_{(j-1)m+1,\ell}^{(\text{in})}}$$

and local average velocity  $\bar{v}_j = m^{-1} \sum_{k=(j-1)m+1}^{jm} v_{k\ell}$ . The local density  $\varrho_j$  is then estimated (as suggested in [2]) via the fluid-dynamic equation

$$\varrho_j = \frac{J_j}{\bar{v}_j}. \quad (2)$$

Such an expression is understood as one of the approximations suitable for estimation of vehicular density. We add that the incorrectness of the definition (2) is caused by the mixing of time and spatial averaging. Besides the macroscopic description of each sample  $S_j$  we now introduce the mean time-clearance (or distance-clearance) by means of definitions

$$\bar{t}_j = \frac{1}{m} \sum_{k=(j-1)m+1}^{jm} t_{k\ell}, \quad \bar{r}_j = \frac{1}{m} \sum_{k=(j-1)m+1}^{jm} r_{k\ell},$$

respectively.

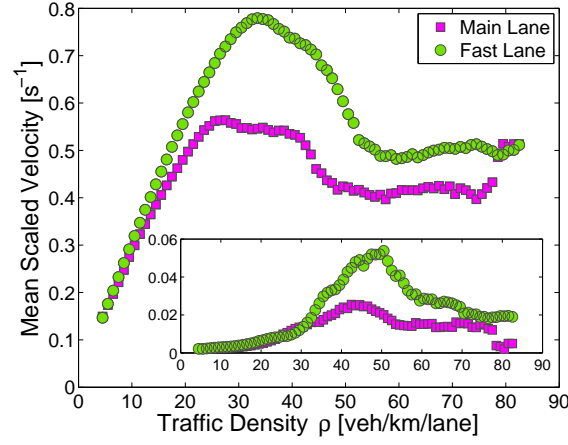


Figure 2: Mean scaled velocity (and variance) as a function of traffic density. In the main part of the figure there is displayed the average value  $\bar{w}$  which is drawn separately for the fast and main lanes. The fluctuations of velocity quantified by the velocity-variance  $\text{VAR}(W) = n^{-1} \sum_{k=1}^n (w_k - \bar{w})^2$  depending on traffic density are visualized in the inset.

Since aiming to investigate the essential properties of micro-distributions (i.e. statistical distributions of microscopic traffic quantities) we eliminate (in the next part of this text) the global trends in those distributions. Predominantly, we eliminate the changes of the average headways caused by the varying traffic density. Such an approach corresponds to the techniques examined in the articles ([8, 11, 9, 21], or [17]), and allows more sophisticated comparison for different traffic regimes or for traffic data originated from different countries. The procedure of the headway re-scaling (see the text below) represents in fact the trivial

variant of the so-called *Savitzky–Golay smoothing filter* (for details please see the Ref. [23]) applied to the matrix spectra in the Random Matrix Theory (see [24]), for example. Thus, we define the *scaled distance-clearances* (for the sample  $S_j$ ) as

$$x_{k\ell} = \frac{r_{k\ell}}{\bar{r}_j} \quad \text{for all } k \in \{(j-1)m+1, (j-1)m+2, \dots, jm\}$$

and the *scaled velocities* as

$$w_{k\ell} = \frac{v_{k\ell}}{\bar{r}_j} \quad \text{for all } k \in \{(j-1)m+1, (j-1)m+2, \dots, jm\}.$$

It implies that the mean clearance in each sample is re-scaled to the unit. However, for completeness of these deliberations we plot (in the Fig. 1) the dependence of the mean distance-clearance on traffic density. Similarly, we analyze the changes (averages and standard deviations) of the scaled velocities  $w_{k\ell}$  in the Fig. 2.

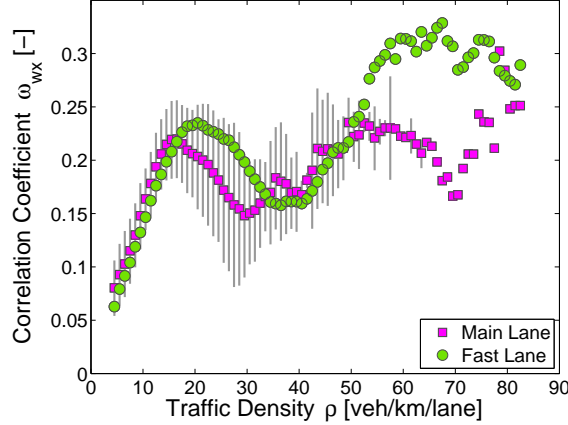


Figure 3: Correlation coefficient as a function of traffic density. The value  $\omega_{wx}$  quantifies a statistical dependence between the scaled vehicular velocities  $w_{k\ell}$  and scaled gaps  $x_{k\ell}$  to the preceding car (in the regions of fixed traffic densities). The squares represent cars moving in the main lane, whereas circles correspond to the fast-lane cars. Note that the vehicles in the fast lane show the stronger correlations than others. Such a tendency is accentuated in the regions where the mental strain of drivers is stronger (i.e. in the regions of over-saturations). Vertical abscissae demonstrate a degree of fluctuations (quantified by the standard deviation) in  $\omega_{wx}$  for main-lane data analyzed from different data sources.

For intentions of analytical calculations executed in one of the next sections it is necessary to explore mutual relations between spatial clearances and car velocities. The standard way how to inspect those relations is to investigate the

correlation coefficient. Denoting  $X = \{x_1, x_2, \dots, x_n\}$  and  $W = \{w_1, w_2, \dots, w_n\}$  the sets of scaled distance-clearances or scaled velocities respectively, one can calculate the Pearson's correlation coefficient using the relation

$$\omega_{wx} = \frac{\text{COV}(X, W)}{\text{VAR}(X) \cdot \text{VAR}(W)} = \frac{\sum_{k=1}^n (x_k - 1)(w_k - \bar{w})}{\sqrt{\sum_{k=1}^n (x_k - 1)^2} \sqrt{\sum_{k=1}^n (w_k - \bar{w})^2}},$$

where  $\bar{w} = \sum_{k=1}^n w_k / n$ . Dividing the complete data-sample into sub-samples (where in the traffic density  $\rho$  is the same, approximately) we demonstrate (in the Fig. 3) that coefficient  $\omega_{wx}$  is localized in the range  $[0.05; 0.30]$  and the corresponding values are depending on a traffic regime (either free or congested regime) and freeway lane. However, in all density-intervals the Pearson's correlation is reduced which means that the hypothesis on a weak dependence between clearance and velocity seems to be legitimate.

### 3. Mathematical derivation of time-clearance distribution

As introduced in [8], justified in [14], and elaborated in [11], one of the possible ways how to acquire meaningful predictions for distributions of empirical distance-clearances is in using the socio-physical traffic model whose elements are repulsed by the short-ranged power-law forces and randomized by the socio-physical noise, which is supposed to be of thermal-like nature. The influence of such a thermodynamical component to the model can be increased/reduced by the socio-physical coefficient  $\beta \in [0, \infty)$  reflecting a mental strain under which the drivers are during a given traffic situation. Specifically, in the free traffic regimes (where the psychological pressure caused by the traffic situation is weak) the parameter  $\beta$  is almost zero. On contrary, for the congested phase (where interactions among drivers are reinforced) the mental strain coefficient  $\beta$  is large. For the detailed changes of  $\beta$  one can inspect the articles [8, 11], or [21].

Specifically, we consider dimensionless particles moving along a ring whose velocities are  $w_1, w_2, \dots, w_N$  and mutual distances (between subsequent particles) are  $x_1, x_2, \dots, x_N$ . Introducing the short-ranged repulsive potential

$$U(x_1, x_2, \dots, x_N) = \sum_{k=1}^N \frac{1}{x_k} \quad (3)$$

and socio-physical hamiltonian (see [11])

$$\mathcal{H}(w_1, w_2, \dots, w_N, x_1, x_2, \dots, x_N) = \frac{1}{2} \sum_{k=1}^N (w_k - w_d)^2 + U(x_1, x_2, \dots, x_N) \quad (4)$$

one can derive (for details please see [11]) that velocities of such a ensemble (analyzed in the steady state) are gaussian-distributed, i.e. the associated probability density reads

$$q(w) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(w-w_d)^2}{2\sigma^2}}, \quad (5)$$

where  $\sigma^2$  is the second statistical moment (variance) and  $w_d$  is the optimal velocity of drivers. Analogously, in the articles [11, 21] there has been deduced that the scaled distance-clearance distribution reads as

$$\wp(r) = A \Theta(r) e^{-\beta/r} e^{-Dr}, \quad (6)$$

where  $\Theta(x)$  is the Heaviside's step-function and the normalization factors are

$$D \approx \beta + \frac{3 - e^{-\sqrt{\beta}}}{2}, \quad (7)$$

$$A^{-1} = 2 \sqrt{\frac{\beta}{D}} \mathcal{K}_1(2\sqrt{D\beta}). \quad (8)$$

As verified in [8, 11, 19] the one-parametric distribution-family (6) is in an satisfactory agreement with the clearance distribution observed in the real-road data. We add that the noise-parameter  $\beta$  is related to the corresponding traffic density  $\varrho$ .

With the help of formulas (5) and (6) one can mathematically derive the analytical form of probability density  $\eta(t)$  for clear time-intervals between two consecutive vehicles (particles of a model). Since the joint probability density for distance and velocity can be (under the condition on weak distance-speed dependency – see the Fig. 3 again) predicted as  $g(w, x) = q(w)\wp(x)$ , the joint probability density for time gaps and velocities reads therefore as  $h(w, t) = wq(w)\wp(wt)$ . Then the time-clearance distribution represents in fact a marginal density

$$\eta(t) = \int_{\mathbb{R}} h(w, t) dw = \int_{\mathbb{R}} wq(w)\wp(wt) dw.$$

After expanding a function  $f(w) = w\wp(wt)$  into the Taylor's series about the optimal velocity  $w_d$  we acquire

$$f(w) = f(w_d) + \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \frac{d^\ell f}{dw^\ell}(w_d)(w - w_d)^\ell,$$



where

$$\frac{d^\ell f}{dw^\ell} = \frac{\partial^\ell \wp}{\partial(wt)^\ell} t^\ell w + \ell \frac{\partial^{\ell-1} \wp}{\partial(wt)^{\ell-1}} t^{\ell-1} \quad (\ell \in \mathbb{N}).$$

Hence

$$\begin{aligned} \eta(t) &= w_d \wp(w_d t) + \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \frac{\partial^\ell \wp}{\partial(wt)^\ell} (w_d t) t^\ell w_d \int_{\mathbb{R}} (w - w_d)^\ell q(v) dw + \\ &+ \sum_{\ell=1}^{\infty} \frac{1}{\ell!} \frac{\partial^{\ell-1} \wp}{\partial(wt)^{\ell-1}} (w_d t) t^{\ell-1} \int_{\mathbb{R}} (w - w_d)^\ell q(w) dw = \\ &= w_d \wp(w_d t) + \sum_{\ell=1}^{\infty} \frac{\mu_\ell}{\ell!} \left( \frac{\partial^\ell \wp}{\partial(wt)^\ell} (w_d t) t^\ell w_d + \frac{\partial^{\ell-1} \wp}{\partial(wt)^{\ell-1}} (w_d t) t^{\ell-1} \right), \end{aligned}$$

where  $\mu_\ell = \int_{\mathbb{R}} (w - w_d)^\ell q(w) dw$  is  $\ell$ -th central statistical moment with three prerogatives cases  $\mu_0 = 1$ ,  $\mu_1 = 0$ , and  $\mu_2 = \sigma^2$ . The latter represents a statistical variance (see also (5)). As well known all the odd central statistical moments (associated to the Gauss distribution) are zero and even moments comply with equalities  $\mu_{2\ell} = \sigma^{2\ell} (2\ell - 1)!!$ , which leads to the general formula

$$\eta(t) = w_d \wp(w_d t) + \sum_{\ell=1}^{\infty} \frac{\sigma^{2\ell}}{(2\ell)!!} \left( \frac{\partial^{2\ell} \wp}{\partial(wt)^{2\ell}} (w_d t) t^{2\ell} w_d + \frac{\partial^{2\ell-1} \wp}{\partial(wt)^{2\ell-1}} (w_d t) t^{2\ell-1} \right). \quad (9)$$

Owing to the facts  $\int_{\mathbb{R}} w_d \wp(w_d t) dt = 1$  and

$$\int_{\mathbb{R}} \left( \frac{\partial^{2\ell} \wp}{\partial(wt)^{2\ell}} (w_d t) t^{2\ell} w_d + \frac{\partial^{2\ell-1} \wp}{\partial(wt)^{2\ell-1}} (w_d t) t^{2\ell-1} \right) dt = 0,$$

the TC-distribution (9) is normalized correctly. In a certain sense (as probably understandable from the two previous equations) the first factor in (9) forms a leading term of TC-distribution, whereas the sum in (9) represents a perturbation term only. Above that, the weight of the  $\ell$ th perturbation summands decreases with a factor  $\sigma^{2\ell}/(2\ell)!!$ . For practical applications (especially for traffic applications) it seems therefore to be meaningful to use an approximate expansion

$$\eta(t) \approx w_d \wp(w_d t) + \frac{\sigma^2}{2} \left( \frac{\partial^2 \wp}{\partial(wt)^2} (w_d t) t^2 w_d + 2 \frac{\partial \wp}{\partial(wt)} (w_d t) t \right). \quad (10)$$

Such an approximation is legitimate because of extremely low value of velocity variance mainly (see the inset in the Fig. 2). Note that the factors  $\sigma^{2\ell}/(2\ell)!!$

measured in real-road data rapidly vanishes since the maximal value of velocity variance is  $\approx 0.02$  in the main lane or  $\approx 0.06$  in the fast lane. The expansion (10) can be re-arranged (after re-scaling clearances) in the formula

$$\eta(t) \approx \wp(t) + \sigma^2 \wp(t) \left( \frac{\beta^2}{2t^2} + \frac{D^2 t^2}{2} - D(t + \beta) \right). \quad (11)$$

The influence of the low velocity-variance to TC-distribution is demonstrated in the Fig. 4 where the curves (11) are compared to the zero-approximation

$$\eta(t) \approx A \Theta(t) e^{-\frac{\beta}{\tau}} e^{-Dt}, \quad (12)$$

where the relations (7),(8) hold true. We remind that the normalization and re-scaling conditions  $\int_{\mathbb{R}} \eta(t) dt = \int_{\mathbb{R}} t \eta(t) dt = 1$  have been applied.

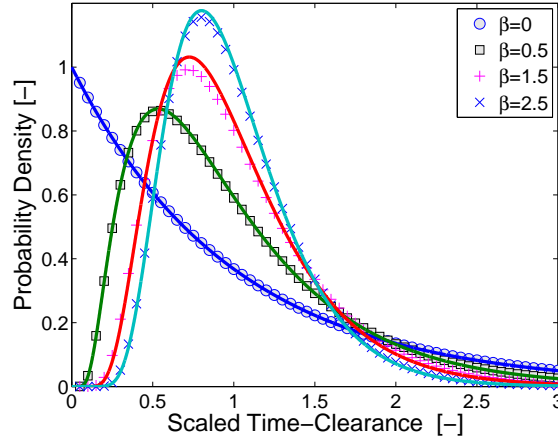


Figure 4: Time-clearance distribution (analytically derived). The signs represent the values of function (11) calculated for  $\beta \in \{0, 1/2, 3/2, 5/2\}$ , whereas the curves display the zero-approximation (12) for the same  $\beta$ . Small discrepancies between signs and corresponding curves are caused by fact that statistical variance  $\sigma^2$  of the scaled velocities (measured on freeways) is very low. Phenomenologically, this relation can be estimated (for the main-lane data) by the inequality  $\sigma^2 \leq (45\beta/55)^4 \exp[-45\beta/14]$ .

#### 4. Mathematical derivation of multi-clearance distribution

Now, knowing the one parameter family of the probability densities (11) or their zero-approximations (12) one can derive an analytical prediction for the so-called *n*th multi-clearance distribution  $\tau_n(t)$  which represents the probability density for a clear time gap  $t$  among  $n + 2$  neighboring particles. Therefore, the main

goal of this section is to quantify (by means of theory of functional convolutions) the probability that time-period between two following instants (the first one: the back bumper of the  $k$ th vehicle is leaving the detector; the second one: the front bumper of the  $(k+n+1)$ th vehicle is intersecting the detector line) is ranging in the interval  $[t, t + dt)$ . Using this notation we find that the probability density for the time-clearance between two succeeding cars (derived in the previous section) is  $\eta(t) = \eta_0(t)$ . Regarding the clearances as independent the  $n$ th probability density  $\eta_n(t)$  can be calculated via recurrent formula

$$\eta_n(t) = \eta_{n-1}(t) \star \eta_0(t),$$

where symbol  $\star$  represents a convolution of the two probabilities, i.e.

$$\eta_n(t) = \int_{\mathbb{R}} \eta_{n-1}(s) \eta_0(t-s) ds.$$

As a first resort, we will concentrate our endeavour to studying the low-density traffic regimes where the locations of cars are independent and therefore the traffic flow corresponds (from mathematical point of view) to the Poisson process of uncorrelated events. As published in [8, 11, 21] such a situation is characterized by the negligible value of socio-physical coefficient  $\beta$  reflecting a driver's mental strain. Thus, the associated clearance distribution reads  $\eta(t) = \Theta(t)e^t$  and sequentially the multi-clearance distribution for free traffic looks like

$$\eta_n(t) = \Theta(t) e^{-t} \int_0^t \int_0^{s_1} \int_0^{s_2} \dots \int_0^{s_{n-2}} \int_0^{s_{n-1}} ds_n ds_{n-1} \dots ds_2 ds_1 = \Theta(t) \frac{t^n}{n!} e^{-t}. \quad (13)$$

As a second resort, we aim to derive a general formula, i.e. formula for non-zero  $\beta$ . For these purposes we use the zero-approximation (12). Applying the method of mathematical induction and an approximation of the function

$$g_n(t, s) = e^{-\beta\left(\frac{n^2}{s} + \frac{1}{t-s}\right)} \approx e^{-\frac{\beta}{t}(n+1)^2} \quad (14)$$

in the saddle point one can obtain

$$\begin{aligned} \eta_n(t) &= \Theta(t) \int_0^t A_{n-1} A s^{n-1} e^{-\beta \frac{n^2}{s}} e^{-Ds} e^{-\frac{\beta}{t-s}} e^{-D(t-s)} ds = \\ &= \Theta(t) A_{n-1} A e^{-Dt} \int_0^t s^{n-1} g_n(t, s) ds \approx \\ &\approx \Theta(t) A_{n-1} A e^{-\frac{\beta}{t}(n+1)^2} e^{-Dt} \int_0^t s^{n-1} ds \approx \Theta(t) A_{n-1} A n^{-1} t^n e^{-\frac{\beta}{t}(n+1)^2} e^{-Dt}. \end{aligned}$$

Hence

$$\eta_n(t) \approx \Theta(t) A_n t^n e^{-\frac{\beta}{t}(n+1)^2} e^{-Dt} \quad (15)$$

where (after applying the re-normalization procedure)

$$A_n^{-1} = 2 \left( \sqrt{\frac{\beta}{D}} (n+1) \right)^{n+1} \mathcal{K}_{n+1}(2(n+1) \sqrt{D\beta}).$$

This fixes the proper normalization  $\int_{\mathbb{R}} \eta_n(t) dt = 1$ . In addition to that the mean  $n$ th spacing equals to

$$\int_{\mathbb{R}} t \eta_n(t) dt = n + 1.$$

Note, that (15) holds true also for the limiting case  $\beta = 0$ . Indeed, with help of formulas  $\lim_{x \rightarrow 0} x^{n+1} \mathcal{K}_{n+1}(x) = (2n)!!$  and  $\lim_{\beta \rightarrow 0_+} D(\beta) = 1$  we easily deduce that

$$\lim_{\beta \rightarrow 0_+} 2 \left( \sqrt{\frac{\beta}{D}} (n+1) \right)^{n+1} \mathcal{K}_{n+1}(2(n+1) \sqrt{D\beta}) = n!.$$

Therefore it holds  $A_n = 1/n!$ , which is in a full consonance with the relation (13). Thus, the relation (15) constitutes a zero-approximation for the distribution of time multi-clearances. Since we have supposed (in the previous deductions) that the variance of scaled velocities is negligible (in a local sense), the form of the distribution (15) is a direct consequence of the distribution for spatial clearances and the fact, that all cars have practically the same velocity (in a local sense, again).

Because the empirical measurements show low (but not negligible) variances in scaled velocities (see the inset in the Fig. 2), it seems more realistic to derive the time multi-clearance distributions under the conditions  $\sigma^2 > 0$  and  $\sigma^{2n} \approx 0$  for  $n = 2, 3, 4, \dots$ . Thus, we suppose that clear time-intervals among succeeding cars are distributed according the rule (11). The detailed analysis of such a relations vindicates that the dominating term in the last summand of (11) is  $\sigma^2 Dt \wp(t)$ . Hence, we surmise that the time clearance follows the low

$$\eta_0^{(\sigma)}(t) \approx \wp(t) - \sigma^2 Dt \wp(t).$$

Here we remark (for mathematical correctness) that this function has to be understood as an approximative probability density, since the perturbation term  $\sigma^2 Dt \wp(t)$  causes that  $\eta_0^{(\sigma)}(t)$  does not fulfil the exact mathematical definition.

Using the method of mathematical induction we prove below that (under the previous surmises) the multi-clearance distribution reads

$$\eta_n^{(\sigma)}(t) \approx \Theta(t) \frac{A^{n+1}}{n!} t^n e^{-(n+1)^2 \frac{\beta}{t}} e^{-Dt} (1 - \sigma^2 Dt). \quad (16)$$

For completeness, we remark that in the following mathematical operations

$$\begin{aligned} \eta_{n+1}^{(\sigma)}(t) &= \eta_n^{(\sigma)}(t) \star \eta_0^{(\sigma)}(t) \approx \Theta(t) \frac{A^{n+2}}{n!} e^{-Dt} \int_0^t e^{-(n+1)^2 \frac{\beta}{s}} s^n e^{-\frac{\beta}{t-s}} ds + \\ &+ D^2 \sigma^4 \Theta(t) \frac{A^{n+2}}{n!} e^{-Dt} \int_0^t e^{-(n+1)^2 \frac{\beta}{s}} (t-s) s^{n+1} e^{-\frac{\beta}{t-s}} ds - \\ &- D \sigma^2 \Theta(t) \frac{A^{n+2}}{n!} e^{-Dt} \int_0^t e^{-(n+1)^2 \frac{\beta}{s}} s^{n+1} e^{-\frac{\beta}{t-s}} ds - \\ &- D \sigma^2 \Theta(t) \frac{A^{n+2}}{n!} e^{-Dt} \int_0^t e^{-(n+1)^2 \frac{\beta}{s}} (t-s) s^n e^{-\frac{\beta}{t-s}} ds = \\ &= \Theta(t) \frac{A^{n+2}}{(n+1)!} t^{n+1} e^{-(n+2)^2 \frac{\beta}{t}} e^{-Dt} (1 - \sigma^2 Dt) \end{aligned}$$

there have been used the approximation (14) and the surmise  $\sigma^4 \approx 0$ .

## 5. Empirical multi-clearance distribution vs. analytical prediction

In this section we will balance the theoretical prognoses deduced in the sections 3 and 4 against the empirical multi-clearance distributions of freeway data analyzed in the section 2. In the following part of the text we will consider the multi-clearances (here: cumulated clearances between five succeeding cars)

$$\kappa_i^{(j)} = \sum_{k=(j-1)m+i}^{(j-1)m+i+3} \frac{t_{k\ell}}{\bar{t}_j}, \quad (\ell = 1, j = 1, 2, \dots, M_1, i = 1, 2, \dots, m-3)$$

enumerated for data sets (1). These multi-clearances are associated with the specific traffic density through the formula (2) and re-scaled so that the average multi-clearance (in the given data-samples  $S_j^{(\text{main})}$ ) is equal to 4. Now, the multi-clearances  $\kappa_i^{(j)}$  represent a statistical realization of random variable  $t$  considered in the section 4 and can be therefore confronted with the theoretical probabilities.

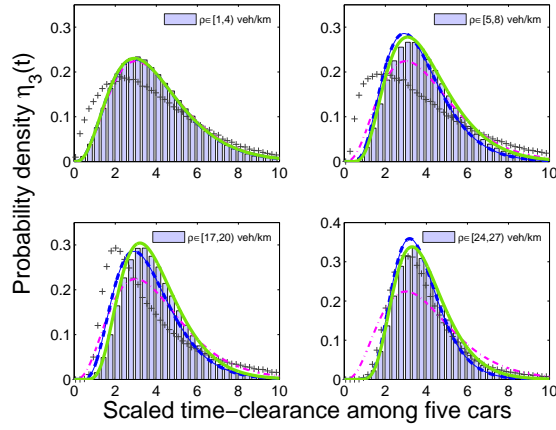


Figure 5: Statistical distribution of time multi-clearances for low density regimes. The bars show the empirical probability density for clear time-interval among five succeeding vehicles moving in main lane (for various density regimes – see legend for details). The blue dashed curves represent the prediction (15) plotted for  $n = 3$  and for the fitted value of coefficient  $\beta = \beta_{\text{fit}}$  obtained by minimizing the weighted error-function (18). Continuous curves (green) display the analytical approximation (17) plotted for  $n = 3$  and for the fitted values of  $\beta = \beta_{\text{fit}}$  and  $\varepsilon = \varepsilon_{\text{fit}}$  specified by minimizing the weighted error-function (19). For clearness, we also plot the dash-dotted curves (magenta) visualizing the multi-clearance distribution (13) valid for an occurrence of independent events/particles. Plus signs illustrate how much the time multi-clearances (gauged by the fast-lane-detectors) differ from those detected in the main lane.

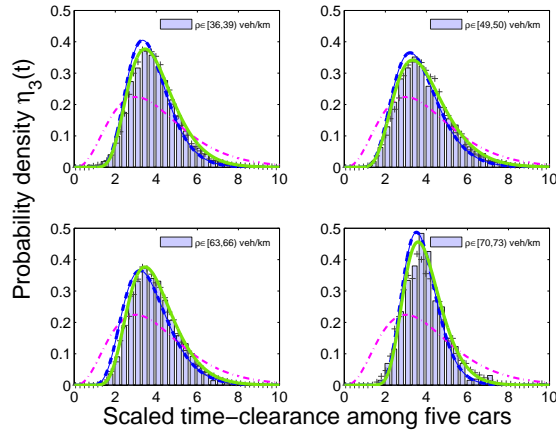


Figure 6: Statistical distribution of time multi-clearances for high density regimes. The bars shows the empirical probability density for clear time-interval among five succeeding vehicles moving in the main lane (for various density regimes – see legend for details). The blue dashed curves represent the prediction (15) plotted for  $n = 3$  and for the fitted value of coefficient  $\beta = \beta_{\text{fit}}$  obtained by minimizing the weighted error-function (18). Continuous curves (green) display the analytical approximation (17) plotted for  $n = 3$  and for the fitted values of  $\beta = \beta_{\text{fit}}$  and  $\varepsilon = \varepsilon_{\text{fit}}$  specified by minimizing the weighted error-function (19). For clearness, we also plot the dash-dotted curves (magenta) visualizing the multi-clearance distribution (13) valid for an occurrence of independent events/particles. Plus signs illustrates how much the time multi-clearances (gauged by the fast-lane-detectors) differ from those detected in the main lane.

To be factual, we divide the entire interval of traffic densities into the subintervals  $[0, 3)$ ,  $[1, 4)$ ,  $[2, 5)$ , and so on and analyze the freeway multi-clearance distribution  $p(\kappa)$  separately in each subinterval. Such an approach has been examined in the articles ([8, 11, 9, 21]) and reflects the known fact that the headway distributions are substantially influenced by the changing location of the system in the phase diagram. To prevent the mixing of the states with different vigilances of car drivers (or with different temperatures of associated heat bath – in the thermodynamic interpretation of vehicular traffic) we make an independent statistical analysis of empirical multi-clearances and investigated their evolution with respect to the traffic density.

In the Fig. 5 and 6 there are plotted the empirical multi-clearance distributions (histograms) against the poissonian distribution (13), zero approximation (15) calculated for  $n = 3$ , and final analytical prediction

$$\eta_3^{(\varepsilon)}(t) = \Theta(t) \frac{A^4}{6} t^3 e^{-16\beta/t} e^{-Dt} (1 - \varepsilon t). \quad (17)$$

The fitted values of the free parameter ( $\beta$  in (15), and  $\beta, \varepsilon$  in (17)) have been determined by means of formulae

$$\beta_{\text{fit}}^* = \operatorname{argmin}_{\beta \in [0, \infty)} \int_0^\infty |\eta_3(t) - p(t)|^2 t e^{-t/4} dt, \quad (18)$$

$$(\beta_{\text{fit}}, \varepsilon_{\text{fit}}) = \operatorname{argmin}_{\beta \in [0, \infty), \varepsilon \in [0, \infty)} \int_0^\infty |\eta_3^{(\varepsilon)}(t) - p(t)|^2 t e^{-t/4} dt, \quad (19)$$

i.e. minimizing the statistical distance  $\chi(\beta, \varepsilon) = \int_0^\infty |f(t) - p(t)|^2 t e^{-t/4} dt$  (weighted by the factor  $t e^{-t/4}$ ) cumulating the weighted deviations between theoretical prediction  $f(t)$  and empirical frequency  $p(t)$ .

As visible in the Fig. 5 and 6 the freeway multi-clearance distribution delineated for low-density states coincides with the distribution (13), which confirms the surmise that vehicles in free traffic regime are moving as independent elements. Thus, their statistics is purely poissonian. As the traffic density rises one can detect larger deviations from (13), which demonstrates the stronger interactions among the cars. Roughly speaking, the fitted thermal parameter  $\beta$  is increasing (in both cases: zero-approximation and also the final probability distribution) with the traffic density (see the Fig. 7). The regions of the temporal descent in  $\beta$  value agree with the critical regions in the fundamental diagram. To be precise,



the transmission between traffic regimes (from free to congested regime) causes the transient consolidation of traffic, which leads to a reduction of driver's mental-strain. Since (see the [21]) the thermal parameter reflects such a level of mental pressure, the detected drop in course of  $\beta = \beta(\varrho)$  is expectable. We add that the evolution of thermal parameter  $\beta$  corresponds to the behavior of the associated quantity investigated within the scope of the articles [8, 11, 21, 26]. Moreover, in all traffic states the distribution (17) fits the real-road data more impressively than original approximation (15), as comprehensible. Such a fact is clearly visible in the Fig. 8 where the statistical distance  $\chi(\beta, \varepsilon)$  between theoretical and empirical distributions are outlined.

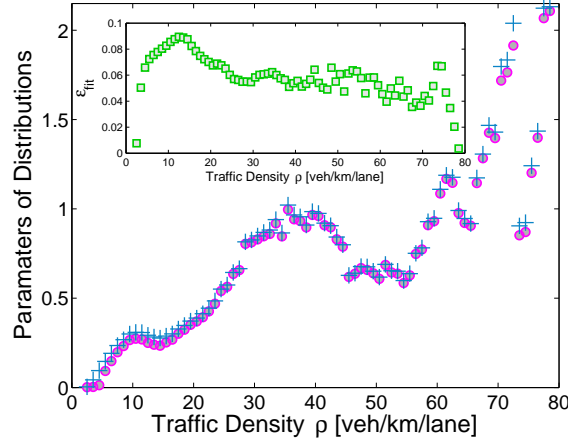


Figure 7: Optimal parameters in theoretical distributions. Circles visualize the value of the fitted parameter  $\beta_{\text{fit}}^*$  (see the relation (18)) that minimizes the statistical distance between theoretical curve (15) and empirical histogram  $p(t)$  (enumerated for main-lane data only). Plus signs correspond to the value of the fitted parameter  $\beta_{\text{fit}}$  in the relation (19) for which the statistical distance  $\chi(\beta, \varepsilon)$  between the theoretical curve (17) and empirical histogram  $p(\kappa)$  is minimal. Similarly, in the inset there is plotted the value of the perturbation parameter  $\varepsilon_{\text{fit}}$  considered in (17).

For illustration purposes (and for better insight into the thermal component of the traffic streams) we also plot the dependency of the thermal parameter  $\beta_{\text{fit}}$  obtained with help of the formula (19) for separate regions inside the fundamental diagram, i.e. we analyze the time multi-clearance distribution in the region with a constant traffic flows and constant traffic density. As comprehensible from the Fig. 9 the basic trends in  $\beta = \beta(\varrho)$  dependency fully corresponds to the behavior discussed above. In the region between 35 and 55 *veh/km/lane* there is

transparently demonstrated how the different traffic regimes (i.e. drivers under the different mental pressures) influence the statistics of vehicular interactions.

For completeness, we emphasize again that all the statistical test discussed in this article have been executed for the main lane traffic data only. As demonstrated in the Fig. 5 and 6 the time multi-clearance distributions detected in the fast-lane data show the well-known discrepancies with the main-lane data (predominantly in the regions of low densities). For denser streams the fast-lane distributions converges to the main-lane distributions, which endorses the belief of scientists that in congested traffic states the correlations among the vehicles in different traffic lanes are much stronger than in free-flow regime.

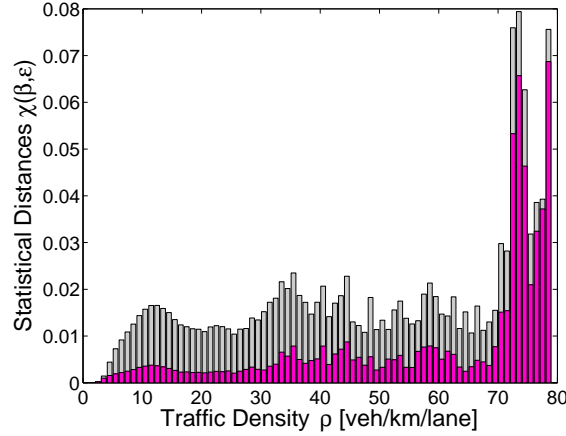


Figure 8: Statistical distance between theoretical and empirical distributions. We plot the values of statistical distances  $\chi(\beta_{\text{fit}}^*)$  and  $\chi(\beta_{\text{fit}}, \varepsilon_{\text{fit}})$  between main-lane multi-clearance distribution  $p(t)$  and zero-approximation (15) (pale bars) or final formula (17) (dark bars), respectively.

## 6. Summary, conclusion, and future prospects

To conclude, this article deals with cumulative time clearances among several subsequent vehicles passing a given point of an expressway. More specifically, we are concentrated predominantly on time intervals between two following occurrences: 1. the rear bumper of  $k$ th car has intersected the detector line, 2. the front bumper of  $(k + n)$ th car has intersected the detector line. Furthermore, we subtract the time-periods when vehicles are occupying the detectors, which means

that a fundamental quantity for this research is the time multi-clearance. Using the local thermodynamical traffic model (originally introduced in [8] and solved analytically in [11]) with short-range repulsive potential among the elements we have derived (applying theory of functional convolutions) a mathematical formula for time multi-clearance distribution. This final theoretical distribution represents an one-parametric family of functions, where the one and only free parameter  $\beta$  reflects the current traffic density. The obtained analytical predictions have been successfully compared with statistics of relevant freeway data. The detected correspondence between theoretical and empirical distributions allows an elaborated insight into changes of vehicular-traffic microstructure influenced by the momentary traffic state.

As perspicuous from the above-mentioned analysis, the statistical distribution of traffic multi-clearances is varying from the Poisson distribution (detected for low-density states) to the low-variance distribution (17) that ascertains a presence of stronger correlations among neighboring vehicles. The lower statistical variance in (17) shows that in the congested traffic states the level of vehicular synchronization is much stronger than in free-flow states. Furthermore, the measure of such a synchronization can be, as discussed in the article, quantified by the thermal-like parameter  $\beta$ .

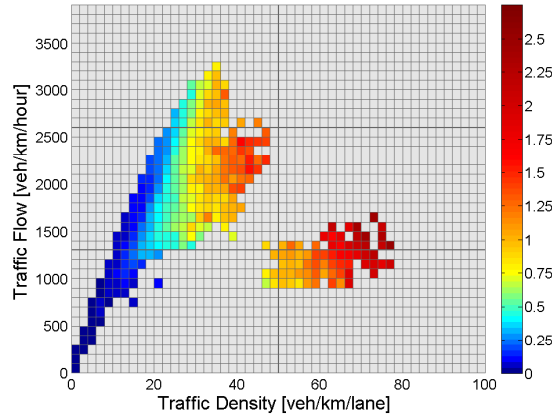


Figure 9: Thermal parameter as a function of traffic density and traffic flux. We plot the dependency of the thermal parameter  $\beta_{\text{fit}}$  occurring in the formula (19) for various regions of the fundamental diagram  $J = J(\rho)$ .

However, the open problem remains how to approximate the time multi-clearance distribution for free-flow vehicles moving in fast lanes. Presence of large amount of vehicular *leaders* in free-traffic regimes, i.e. comparable percentage of leaders and *followers*, causes that the clearance distribution is probably a compound of two partial distributions (first one for leaders, second one for followers). Thus, the expected approach leading to analytical predictions for fast-lane clearances can be found in the theory of finite mixture distributions (similarly to semi-poisson model discussed in [5]).

#### *Acknowledgement*

The author would like to thank Cecile Appert-Rolland (Laboratoire de Physique Théorique Université de Paris-Sud, Orsay) for valuable remarks which have been conducive to the presented research. This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the project MSM 6840770039 and by the Czech Technical University within the project SGS12/197/OHK4/3T/14.

#### **References**

- [1] D. Chowdhury, L. Santen, and A. Schadschneider, *Physics Reports* **329** (2000) 199
- [2] D. Helbing, *Rev. Mod. Phys.* **73** (2001) 1067
- [3] S. Hoogendoorn and P. Bovy, *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering* **215/4** (2001) 283
- [4] B.S. Kerner, *The Physics of Traffic*, Berlin, New York: Springer Verlag (2004)
- [5] D.J. Buckley, *Transportation Science* **2/2** (1968) 107
- [6] M. Krbálek, P. Šeba, and P. Wagner, *Phys. Rev. E* **64** (2001) 066119
- [7] W. Knospe, L. Santen, A. Schadschneider, and M. Schreckenberg, *Phys. Rev. E* **65** (2002) 056133
- [8] M. Krbálek and D. Helbing, *Physica A* **333** (2004) 370
- [9] D. Helbing, M. Treiber, and A. Kesting, *Physica A* **363** (2006) 62

- [10] M. Treiber, A. Kesting, and D. Helbing, Phys. Rev. E **74** (2006) 016123
- [11] M. Krbálek, J. Phys. A: Math. Theor. **40** (2007) 5813
- [12] A.Y. Abul-Magd, Phys. Rev. E **76** (2007) 057101
- [13] A. Šurda, J. Stat. Mech., **04** (2008), P04017
- [14] M. Treiber and D. Helbing, Eur. Phys. J. B **68** (2009) 607
- [15] L. Li, F. Wang, R. Jiang, J. Hu, and Y. Ji, Chinese Phys. B **19** (2010) 020513
- [16] X. Chen, L. Li, R. Jiang, and X. Yang, Chinese Phys. Lett. **27** (2010) 074501
- [17] M. Krbálek and P. Hrabák, J. Phys. A: Math. Theor. **44** (2011) 175203
- [18] D. Helbing and M. Treiber, Phys. Rev. E **68** (2003) 067101
- [19] M. Krbálek, J. Phys. A: Math. Theor. **41** (2008) 205004
- [20] C. Appert-Rolland, Phys. Rev. E **80** (2009) 036102
- [21] M. Krbálek and P. Šeba, J. Phys. A: Math. Theor. **42** (2009) 345001
- [22] X. Jin, Y. Zhang, F. Wang, L. Li, D. Yao, Y. Su, and Z. Wei, Transportation Research Part C: Emerging Technologies, **17/3** (2009) 318
- [23] A. Savitzky and M.J.E. Golay, Analytical Chemistry **36/8** (1964) 1627
- [24] M.L. Mehta, *Random matrices (Third Edition)*, New York: Academic Press (2004)
- [25] A. Sopasakis, Physica A **342** (2004), 741
- [26] M. Krbálek, Kybernetika **46/6** (2010), 1108